Online Appendix for Violence and Election Fraud: Evidence from Afghanistan

Nils B. Weidmann
Centre for the Study of Civil War
Peace Research Institute Oslo
0186 Oslo, Norway
nils.weidmann@gmail.com

Michael Callen
Department of Economics
University of California, San Diego
La Jolla, CA 92093, USA
mjcallen@ucsd.edu

December 2011

1 The IEC Recount

Initially, the ECC designated stations in the following categories as eligible for inspection:

- **A1** - Polling stations in which 600 or more valid votes were cast
- **B1** - Polling stations with more than 100 votes in which one candidate received 95 percent or more of the total votes cast
- **C1** - Polling stations in which 600 or more valid votes were cas AND in which one candidate received 95% or more of the total votes cast

Miscommunication between the IEC and the ECC resulted in the initial sample deviating from the terms of the initial order in two ways. First, the initial sample was drawn on the basis of valid votes rather than total ballots, valid and invalid, cast. Second, the initial sample included stations that had been quarantined because of complaints filed with the ECC. According to the ECC's interpretation of the initial order, quarantined stations should not be eligible for sampling during the recount. Consequently, three additional categories of ballot boxes were designated as eligible for the recount. These categories excluded stations already covered by the previous categories.

- **A2** - Polling stations in which 600 or more total votes were cast
- **B2** - Polling stations in which one candidate received 95 percent or more of the total valid votes cast
- **C2** - Polling stations in which 600 or more total votes were cast AND in which one candidate received 95% or more of the valid votes

The 95% of total votes requirement, in effect, means that the ballots in Category C2 are a strict subset of those in Category B1. 3,369 polling stations qualified as potentially fraudulent based on these six criteria. In order to quickly estimate the extent of fraud, the IEC and ECC agreed to draw a 10% sample (342 boxes) of the 3,369 boxes for physical inspection. The rationale for withdrawing boxes with 600 votes were that 600 represents 100 percent turnout

based on the IEC's pre-poll estimate of the voting population. The key signs of manipulation that were checked during the physical inspection were whether: (a) the box was intact, sealed, and appropriately marked; (b) the reconciliation forms and result forms were included in the box; (c) there existed some discrepancy between total ballots bundled and total votes recorded; (d) ballot papers appeared to be folded individually by voters; and (e) whether there existed significant patterns of similar markings.[1] The data provided by this recount allow us to construct an estimate of fraud at the district level and to directly test the performance of digit-based forensic fraud measures.

[1] See "Audits at a Glance", accessed on July 4, 2010

2 Comparing Fraud Measures

As an additional comparison between the recount-based and last-digit fraud measures, we estimate regression models predicting the recount-based fraud measure using our last-digit test. In the first model, we apply this test at the district level, using the measures described in the main text. Second, we investigate whether the last-digit measure, applied at the district level, predicts fraud in a polling-station level regression on just the audited sample of 342 boxes. Both tests provide evidence that the chi-squared measure proposed by Beber and Scacco provides a valid measure of ballot box stuffing.

	Model A1
Fraud, Total Count, χ^2	0.325***
	(0.059)
Constant	0.072***
	(0.010)
R-squared	0.357
N	389

Table 1: Comparing fraud measures at the district level. Dependent variable: estimated fraud share. Level of significance: $*p < 0.1$, $**p < 0.05$, $***p < 0.01$. Robust standard errors clustered at the province level reported in parentheses.

We see that all the last-digit measure strongly predicts fraud, accounting for more than 35% of the variance in our recount-based measure. The recount-based and digit-based fraud measures are fundamentally different in their construction. Finding a clear correlation gives us confidence that both measures effectively measure ballot fraud. However, we can perform a much more powerful test by checking whether the last-digit measure predicts fraud in the recount sample itself, which is what we the next regression model does. Here, our dependent variable is a dummy equal to 1 if polling station i in district d is found to be fraudulent through physical inspection. We predict fraud as identified through physical inspection by the last-digit indicator computed on all polling stations in the respective district. We include dummies for the ECC categories (see above) to allay concern that the Chi-square test is only picking up boxes brought into the sample because they record 600, the maximum allowable number of votes for any station. Detecting fraud on this sample is difficult. The 342 polling stations were drawn explicitly based on the probability that they are fraudulent and so have very limited variation in fraud. 73.1% of the polling stations in this sample exhibit physical signs of fraud.

Again, we find that the chi-squared digit based measure predicts fraud on the recount sample, although the explained variance is much lower than above. Still, together these results provide evidence that both the recount-based measure and the Beber-Scacco chi-squared test applied at the district level successfully measure some of the ballot box manipulation that took place during Afghanistan's 2009 election.

	Model A2
Fraud, Total Count, χ^2	0.180***
	(0.051)
Constant	0.796***
	(0.050)
R-squared	0.130
N	342

Table 2: Testing the last-digit measure on the audited sample. Dependent variable: verified fraudulent in recount. Levels of significance: $^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. All regressions include dummies for ECC categories: A1 - 600 or more valid votes cast; A2 - 600 or more valid and invalid votes cast; B1 - 95% or more valid and invalid votes for one candidate and; B2 - 95% or more valid votes cast for one candidate. Robust standard errors clustered at the district level reported in parentheses.

3 Additional Tests for the Inverted-U Shape

The following table reports the results of additional test for the inverted-U shaped relationship between violence and fraud, using the ACLED[2] and WITS[3] event datasets. In addition to the fraud measures used in the paper (last-digit test and recount-based estimate), we include simpler indicators used by the ECC and the IEC (share of stations with 600+ votes, share of stations with 95% or more for one candidate).

ACLED data				
Dependent Variable	Shr of Stations w/ 95% for One Cand	Shr of Stations w/ 600+ Votes	Total Count (χ^2)	Recount Fraud Share
Events (ACLED)	0.204**	0.021	0.552*	0.088
	(0.093)	(0.046)	(0.295)	(0.231)
Events (ACLED)^2	-0.177**	-0.033	-0.494**	-0.120
	(0.071)	(0.036)	(0.233)	(0.174)
R-squared	0.178	0.116	0.214	0.250
N	375	375	372	372
WITS data				
Events (WITS)	0.085**	0.049**	0.291***	0.148*
	(0.039)	(0.020)	(0.101)	(0.079)
Events (WITS)^2	-0.009	-0.011	-0.079*	-0.031
	(0.022)	(0.010)	(0.040)	(0.028)
R-squared	0.193	0.137	0.227	0.277
N	375	375	372	372

Table 3: Level of significance: $^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. All regressions control for the number of stations closed, per capita expenditure, and land inequality. Robust standard errors clustered at the province level reported in parentheses.

[2]Clionadh Raleigh et al. "Introducing ACLED: An Armed Conflict Location and Event Dataset". In: *Journal of Peace Research* 47.5 (2010), pp. 651–660.

[3]National Counterterrorism Center. *Worldwide Incidents Tracking System (WITS).* Available at http://wits.nctc.gov/. 2010.

4 Is Violence Perpetrated to Strategically Favor a Given Candidate?

Figure 1 depicts the relationship between ethnolinguistic fractionalization (ELF) and violence around election day across districts in Afghanistan. As voting is broadly along ethnic lines in Afghanistan, we should expect that violence is increasing in ELF if, to some degree, it is targeted to favor a given candidate. However, there does not appear to be any relation between ethnolinguistic fractionalization and violence in our data.

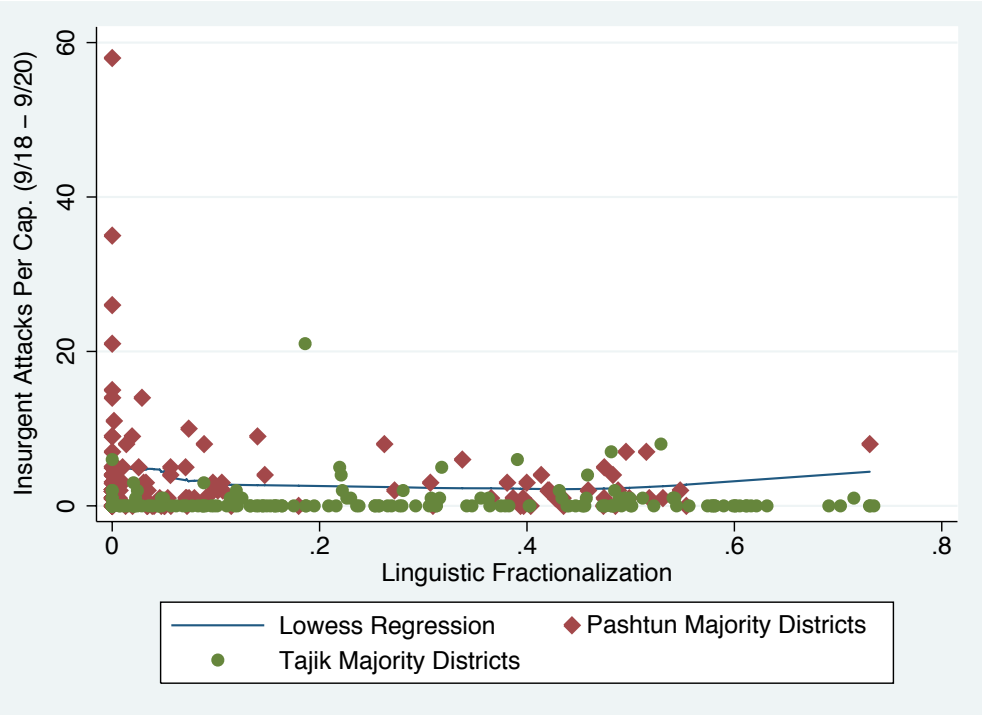

Figure 1: Violent incidents per capita from August 18 - August 20, 2009 and ethnolinguistic fractionalization